Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

CRT-2003-12

CONSTRAINED ESTIMATION OF A ROAD SAFETY COVARIANCE MATRIX USING SCHUR COMPLEMENTS

by

Assi N'Guessan

May 2003

Université de Montréal
C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7 Canada
Téléphone : (514) 343-7575 / Télécopieur : (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

HEC ÉCOLE POLYTECHNIQUE MONTRÉAL Université de Montréal

103-07103/

PUBLICATION
CRT-2003-12

CONSTRAINED ESTIMATION OF A ROAD SAFETY COVARIANCE MATRIX USING SCHUR COMPLEMENTS

by

Assi N'Guessan[1]

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

1. École Polytechnique Universitaire de Lille et Laboratoire de Mathématiques Appliquées, CNRS, FRE 2222. Université de Lille 1, 59655 Villeneuve d'Ascq Cedex, France (Assi.Nguessan@polytech-lille.fr)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2003

Abstract

One is interested in the asymptotic matrix of the constrained m.l.e. (maximum likelihood estimator) of a multidimensional parameter connected to a road safety measure. He achieves the explicit asymptotic covariance matrix of the constrained m.l.e. and the explicit form of certain estimated standard errors through Schur complements technical inversion method. This approach also provides a powerful tool to compute formally the explicit estimated between and within variances covariance's matrices of block-components of the restraint m.l.e. ease of implementation without matrix program inversion is an attractive aspect of this new approach. Some examples of formal estimation are discussed to back up the method.

Keywords: Road safety measure, multinomial model, constrained maximum likelihood, Schur complements, asymptotic covariance's matrix, formal estimation.

Résumé

On s'intéresse à la matrice de covariances asymptotique du m.l.e. (estimateur du maximum de vraisemblance) sous contraintes d'un paramètre multidimensionnel lié à une mesure de sécurité routière. Il fait appel à une technique formelle d'inversion matricielle appelée compléments de Schur pour obtenir explicitement la matrice de covariances asymptotique du m.l.e. sous contraintes ainsi que l'estimation formelle de certains écart-types. Cette approche est un outil puissant qui permet également d'estimer la structure analytique des matrices de variances covariances inter et intra de bloc d'éléments du m.l.e. sous contraintes. La mise en œuvre simple et sans obligation de programmes informatiques d'inversion matricielle de cette nouvelle approche le rend d'autant plus intéressante. Quelques exemples d'estimation formelle sont fournis pour étayer la méthode.

Mots-clés: mesure de sécurité routière, modèle multinomial, maximum de vraisemblance sous contraintes, compléments de Schur, matrice de covariances asymptotique, estimation formelle.

1. Introduction and review

In many estimation problems, particulary in multivariate statistics, the parameters in a model are not functionally independent, that is, there exist certain relations connecting them. Nevertheless, it may be undesirable to eliminate redundancies for reasons of symmetry, or because the natural interpretation of the parameters would be lost, or because of mathematical untractability. Statistical aspects of estimation in restricted parameter spaces, asymptotic results for constrained estimation and related problems have been extensively discussed (*see* for example Harville 1997, Lee 1979, Van Eeden 1996).

More precisely, Aitchison and Silvey (1958) consider the case of i.i.d. observations. They prove, under certain conditions, asymptotic existence, consistency and asymptotic normality of a solution of the likelihood equations. Silvey (1959) deals with the i.i.d. case, but allowing the information matrix to be singular. He uses the alternative approach to proof of consistency due to Wald (1949). Crowder (1984) considers the case of non i.i.d. observations when the parameters are constrained and uses inequality constraints as an application. His asymptotic results and formal methods stem from those of Aitchison and Silvey (1958) and Silvey (1959). Also, more emphasis is placed in Crowder's (1984) paper on the precise manner in which the constraints may be used to construct a suitable nonsingular version of the information matrix and the connection with parameter identifiability. Neuenschwander and Flury (1997) give an alternative proof of a theorem by Silvey (1959) for the asymptotic covariance matrix of the maximum likelihood estimation (m.l.e.) under different types of constraints in the parameter space. They show how the different types of constraints can be treated in a unified way and applied to a variety of statistical estimation problems.

N'Guessan et al (2001) deal with the constrained estimation of the average effect of a road safety measure using before-after accident data tables. The latter authors consider independent non i.d. multinomial observations under linear constraints and are interested in the existence and convergence of the restraint m.l.e. connected to their model via simulation studies and mean square error. But they don't include the accuracy of the constrained m.l.e. i.e. the asymptotic covariances matrix. Also, their method's drawback is that both parameters's space dimension and constraints's number increase with the size of before-after data tables used and thereafter the dimension of Fisher's information matrix to invert. One knows that this situation sometimes affects the numerical inversion results and impair the accuracy of standard errors.

The main motivation here is to make their paper complete by investigating the asymptotic variances covariances matrix of their m.l.e. in order to obtain an explicit form of estimated standard errors without calling numerical matrix inversion programm. The formal methods combine Fisher's information matrix related to model below, the notion of Schur complements for the inversion of partitioned matrix (Ouellette 1981; N'Guessan 2003a) and the results of Crowder (1984).

The constrained multinomial model and notation in section 2, used to share the total accident number in each experimental site at the same time between both periods and among the different accident types, stem from those of N'Guessan et al (2001). In section 3 some techinical intermediate results are derived to support the foregoing theoren and its corollaries. We particulary use the model and constraints below to derive that the Fisher's information matrix is nonsingular. Section 4 is exclusively centred on the main result. its proof is given in the appendix. In section 5 some formal estimation examples are discussed to back up the method . Some concluding remarks are displayed in section 6.

2. Statistical Model and notation

The statistical models used in the before-after estimation of a road safety measure efficiency heavily depend on the available data and on the targets set along with the measure(*see* for example Danielson (1986 , 1988), Hauer (1997) , Barry et al (1999), Abdel-Aty et al (2000), Gaudry and Lassarre (2000), Beenstock et al (2001). Some factors such as traffic evolution or weather conditions, etc... must be considered when interpreting the estimation results; if when studying other data (such as the variations of the accident number on a similar site where the measure is not applied) one can notice that the accident number is multiplied by a c factor (control or evolution coefficient), that factor must be taken into account when appreciating the measure's efficiency. When a measure is simultaneously applied to several sites, considering the use of that very factor whatever the site is generally not realistic because of site location for example. That is why Tanner (1958) introduced the notion of control area and defined a control coefficient for each site using the ratio of the accident number after to before in the control area. Let's now suppose that, the measure having been applied to several sites, the total number of accidents for both periods in each target site is avalaible. Then let's retrospectively focus on the measure's average effect in each site on a certain number of accidents (fatal accidents, seriously injured people, slightly injured people, material damage, etc...). Such an approach introduces a third dimension (accident typology) to Tanner's pattern and statistics and imposes the estimation of the measure's average effect through before-after tables of accident data. With that in mind, N'Guessan (1993), N'Guessan and Langrand (1993) generalised Tanner's approach to before-after tables of accident with control area. In their papers they supposed that the same road safety measure had been applied to several sites, each one including mutually exclusive different accident types and linked to a specific control area. Recently, N'Guessan et al (2001) are interested in the constrained m.l.e. of the average effect of a road safety measure in presence of random control area data and focus their attention on its numerical convergence via the value of the mean square error for before-after tables. They deal with the multidimensional estimation of the average effect of road safety measure applied to $s(s > 0)$ entities (sites), each one including $r(r > 1)$ mutually exclusive different accident types and linked to a specific control area. They consider a collection

$$(Y_{11}, Y_{21}), (Y_{12}, Y_{22}), \ldots, (Y_{1s}, Y_{2s})$$

of s independent couples of $2r-$ dimensional random vector and assume that , for fixed k, each couple (Y_{1k}, Y_{2k}) has multinomial distribution denoted by $\mathcal{M}(n_k; \Pi_{1k}(\alpha), \Pi_{2k}(\alpha))$, where

$$\Pi_{tk}(\alpha) = (\pi_{t1k}(\alpha), \pi_{t2k}(\alpha), \ldots, \pi_{trk}(\alpha)),$$

$n_k(n_k > 0)$ is a given integer, $\alpha = (\beta_0, \beta^T)^T \in \mathbb{R}^{1+sr}$ a unknown vector parameter, $\beta_0(\beta_0 > 0)$ a scalar, $\beta = (\beta_1^T, \ldots, \beta_s^T)^T \in \mathbb{R}^{sr}$ with $\beta_k = (\beta_{1k}, \beta_{2k}, \ldots, \beta_{rk})^T \in \mathbb{R}^r$ and $\beta_{jk} > 0$. The components $\pi_{trk}(\alpha)$ of the $r \times 1$ vector probability $\Pi_{tk}(\alpha)$ are given by a continuous differentiable functional link $g : \mathbb{R}^{1+sr} \mapsto]0, 1[$ so that $\pi_{trk}(\alpha) = g(z_k, \alpha)$ and

$$g(z_k, \alpha) = \begin{cases} \dfrac{\beta_{jk}}{1 + \beta_0 < z_k, \beta_k >} & \text{t=1; j=1,2,...,r} \\ \\ \dfrac{z_{jk}\beta_0\beta_{jk}}{1 + \beta_0 < z_k, \beta_k >} & \text{t=2; j=1,2,...,r} \end{cases} \tag{1}$$

where $<,>$ is the usual inner product, $z_k = (z_{1k}, z_{2k}, \ldots, z_{rk})^T$ a given extra set data connected to the control area of each couple k. The latter authors are interested in the m.l.e. $\hat{\alpha}$ of α and its numerical consistency under the restraints $h_k(\alpha) = 0$ where

$$h_k(\alpha) = < \mathbf{1}_r, \beta_k > -1; k = 1, 2, \ldots, s \tag{2}$$

with $\mathbf{1}_r = (1,\ldots,1)^T \in \mathbb{R}^r$, the vector of unity. The construction mechanism of the statistical patter above is displayed in N'Guessan et al (2001). Let's indicate that in road safety modeling β_0 represents the average effect when combining all data tables and vector β_k , for fixed k , denotes the probability distribution (or risk accident) connected to extra data z_k of control area k such that $\sum_{j=1}^{r} \beta_{jk} = 1$. This is the natural and practical interpretation of the constraints (2).

In the following we note $H_\alpha = (H_1, H_2)$ the constraints matrix with $H_1 = \frac{\partial h}{\partial \beta_0}$ a $s \times 1$ matrix, $H_2 = \frac{\partial h}{\partial \beta}$ a $s \times (sr)$ matrix, $h = (h_1, h_2, \ldots, h_s)$, $L(\alpha)$ the log-likelihood function based on the s couples, $J_\alpha = E(-\frac{\partial^2 L}{\partial \alpha \partial \alpha})$, the $(1+sr) \times (1+sr)$ information matrix evaluated at different points on the line segment joining α^0 (the true vector parameter) and $\hat{\alpha}$. It can been shown under regularity conditions, that $\hat{\alpha}$ is asymptotically normally distributed and its covariance matrix strongly depends on the fact that both J_α and $H_\alpha J_\alpha^{-1} H_\alpha^T$ are nonsingular. Technical lemmas below elaborate these points. The foregoing theorem concerns the main result and gives the formal expressions of the components of the asymptotic covariance matrix connected to model (1) and constraints (2). We achieve those explicit forms through the notion of Schur complements well presented by Ouellette (1981). She also presents an unified traitment of standard results on inverses , and g-inverses of partitioned matrices using Schur complements and discribes some areas of statistics in which this notion arises. It is the same for spatial prediction and temporal backcasting (Le et al. 2001) or log-linear model estimators (Rivest and Lévesque, 2001) or longitudinal data models (Frees, 2001) and so on.

3. Technical intermediate results

lemma 1 *Let's note* $V_{\alpha,k} = \frac{\beta_0}{1+\beta_0 < z_k, \beta_k >}(z_{1k}, z_{2k}, \ldots, z_{rk})^T$ *a* $r \times 1$ *vector and* $\Omega_{\alpha,k} = diag(\frac{1+\beta_0 z_{1k}}{\beta_{1k}}, \frac{1+\beta_0 z_{2k}}{\beta_{2k}}, \ldots, \frac{1+\beta_0 z_{rk}}{\beta_{rk}})$ *a* $r \times r$ *matrix. Then for* $k = 1, 2, \ldots, s$*:*
i) $\| z_k \|^2_{\Omega^{-1}_{\alpha,k}} < \frac{< z_k, \beta_k >}{\beta_0}$; *ii)* $\| V_{\alpha,k} \|^2_{\Omega^{-1}_{\alpha,k}} < \frac{\beta_0 < z_k, \beta_k >}{(1+\beta_0 < z_k, \beta_k >)^2} < 1$.

Proof. i) Using the fact that $\frac{x}{1+x} < 1$ for all $x > 0$, one obtains after matrix manipulations, $\| z_k \|^2_{\Omega^{-1}_{\alpha,k}} < \frac{1}{\beta_0}(\sum_{m=1}^{r} \beta_{mk} z_{mk}) = \frac{< z_k, \beta_k >}{\beta_0}$.
ii) This part follows from i) and the relation between the vectors V_k and z_k ■

lemma 2 *Let's note* $B_{\alpha,k} = \gamma_k(\Omega_{\alpha,k} - V_{\alpha,k} V^T_{\alpha,k})$ *with* $\gamma_k = \frac{n_k}{1+\beta_0 < z_k, \beta_k >}$, *Then*
i) For $k = 1, 2, \ldots, s$;
$B^{-1}_{\alpha,k} = \gamma_k^{-1}(\Omega^{-1}_{\alpha,k} + t_k \Omega^{-1}_{\alpha,k} V_{\alpha,k} V^T_{\alpha,k} \Omega^{-1}_{\alpha,k})$; $t_k = (1 - \| V_{\alpha,k} \|^2_{\Omega^{-1}_{\alpha,k}})^{-1}$
ii) $\| U_\alpha \|^2_{B^{-1}_\alpha} < \tau_\alpha$; *where* $B_\alpha = Bloc-diag(B_{\alpha,1}, \ldots, B_{\alpha,s})$, $U_\alpha = (U^T_{\alpha,1}, \ldots, U^T_{\alpha,s})^T$ *a* $(sr) \times 1$ *vector with* $U_{\alpha,k} = \frac{\gamma_k}{\beta_0} V_{\alpha,k}$ *and* $\tau_\alpha = \sum_{k=1}^{s} \frac{\gamma_k < z_k, \beta_k >}{\beta_0 (1+\beta_0 < z_k, \beta_k >)}$.

Proof. i) Lemma 1 implies that $t_k > 0$ for all k, so $B^{-1}_{\alpha,k}$ follows from Woodburry (1950) formula (*see* also Press et al (1992) pp. 73-78 , Ouellette 1981) or Sherman-Morrison formula (1950).
ii) Matrix manipulations and relations on the one hand between vectors $U_{\alpha,k}$ and $V_{\alpha,k}$ and on the other hand matrices $B_{\alpha,k}$ and $\Omega_{\alpha,k}$ imply that

$$\| U_\alpha \|^2_{B^{-1}_\alpha} = \frac{1}{\beta_0^2} \sum_{k=1}^{s} \frac{\gamma_k \| V_{\alpha,k} \|^2_{\Omega^{-1}_{\alpha,k}}}{1 - \| V_{\alpha,k} \|^2_{\Omega^{-1}_{\alpha,k}}} \tag{3}$$

Using lemma 1, one obtains

$$\| U_\alpha \|^2_{B_\alpha^{-1}} < \sum_{k=1}^{s} \frac{\gamma_k(1+\beta_0 < z_k, \beta_k >)}{\beta_0^2} \| V_{\alpha,k} \|^2_{\Omega^{-1}_{\alpha,k}} < \tau_\alpha \tag{4}$$

■

Now we are ready to deal with the inversion of J_α and $H_\alpha J_\alpha^{-1} H_\alpha^T$. Using model (1) above and after some derivations and matrix manipulations one obtains:

$$J_\alpha = \begin{pmatrix} \tau_\alpha & U_\alpha^T \\ U_\alpha & B_\alpha \end{pmatrix}.$$

Following Ouellette (1981), and N'Guessan (2003a, appendix B) we note (J_α/B_α) (resp. (J_α/τ_α)) the Schur complement of B_α (resp. τ_α) in J_α and defined by $(J_\alpha/B_\alpha) = \tau_\alpha - U_\alpha^T B_\alpha^{-1} U_\alpha$ (resp. $(J_\alpha/\tau_\alpha) = B_\alpha - U_\alpha \tau_\alpha^{-1} U_\alpha^T$).

lemma 3
(i) $(J_\alpha/B_\alpha) > 0$
(ii) (J_α/τ_α) *is nonsingular and*

$$(J_\alpha/\tau_\alpha)^{-1} = B_\alpha^{-1} + (J_\alpha/B_\alpha)^{-1} B_\alpha^{-1} U_\alpha U_\alpha^T B_\alpha^{-1} \tag{5}$$

Proof. i) is a consequence of lemma 2 part ii)
ii) This result is derived from the combination of lemma 2 and Woodburry formula a special case of Schur complements inversion. ■

lemma 4
i) J_α^{-1} *exits and*

$$J_\alpha^{-1} = \begin{pmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1} U_\alpha^T B_\alpha^{-1} \\ -B_\alpha^{-1} U_\alpha (J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{pmatrix} \tag{6}$$

ii)

$$(H_\alpha J_\alpha^{-1} H_\alpha^T)^{-1} = \Lambda_\alpha^{-1} - ((J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}})^{-1} \Lambda_\alpha^{-1} \xi_\alpha \xi_\alpha^T \Lambda_\alpha^{-1} \tag{7}$$

Proof. i) uses lemma 2 and the formal inversion of partitioned matrix via Shur complements.
ii) combines the constraints functions (2) and i). More precisely, straightforward derivations of constraints (2) show that

$$\frac{\partial h_k}{\partial \beta_m} \equiv \mathbf{1}_r^T \delta_{mk}; \qquad k = 1, 2, \ldots, s; m = 0, 1, 2, \ldots, s \tag{8}$$

where δ_{mk} is the Kronecker symbol. So using the latter formula , part i) of lemma 4 and ii) of lemma 3 we can write:

$$\begin{aligned} H_\alpha J_\alpha^{-1} H_\alpha^T &= H_2 (J_\alpha/\tau_\alpha)^{-1} H_2^T - H_1 (J_\alpha/B_\alpha)^{-1} \xi_\alpha^T \\ &= H_2 (J_\alpha/\tau_\alpha)^{-1} H_2^T \\ &= \Lambda_\alpha + (J_\alpha/B_\alpha)^{-1} \xi_\alpha \xi_\alpha^T \end{aligned} \tag{9}$$

The same arguments used for part i) of lemma 2 are applied. ■

4. Main result

The restraints m.l.e of α_0 , provided it exists, is the vector $\hat{\alpha}$ which satisfies

$$(\frac{\partial L}{\partial \alpha})_{\hat{\alpha}} + H_{\hat{\alpha}}^T \hat{\lambda} = 0, \quad h(\hat{\alpha}) = 0 \tag{10}$$

where λ is a vector of s Lagrange multipliers. Let $M_\alpha = J_\alpha^{-1}\left(J_\alpha + (\frac{\partial^2 L}{\partial \alpha \partial \alpha})\right)$ a $(1+sr)\times(1+sr)$ matrix , $R_\alpha = H_2(J_\alpha/\tau_\alpha)^{-1}H_2^T$ a $s\times s$ matrix , $\Phi_\alpha = J_\alpha^{-1}H_\alpha^T R_\alpha^{-1}$ a $(1+sr)\times s$ matrix , $\xi_\alpha = H_2 B_\alpha^{-1}U_\alpha$ a $s\times 1$ vector, $\Lambda_\alpha = diag(\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,1}}, \parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,2}}, \ldots, \parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,s}})$ a $s\times s$ matrix with $\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,k}} = \mathbf{1}_r^T B^{-1}_{\alpha,k}\mathbf{1}_r$.

Theorem 4.1 *Suppose that*
Assumption A1 : $(\hat{\alpha}, \hat{\lambda})$ *is a consistent solution of the latter equation*
Assumption A2 : $M_{\hat{\alpha}} \overset{Proba}{\longrightarrow} 0$, $\Phi_{\hat{\alpha}} - \Phi_{\alpha^0} \overset{Proba}{\longrightarrow} 0$,$\mid \Phi_{\alpha^0} \mid < \infty$.
Assumption A3 : $(\frac{\partial L}{\partial \alpha})_{\alpha^0} \overset{Distr.}{\sim} \mathcal{N}(0, J_{\alpha^0})$
Then the asymptotic components of the covariance matrix of $\hat{\alpha}$ *are:*
$\sigma^2(\hat{\beta}_0) = (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-2} \parallel \xi_\alpha \parallel^2_{R_\alpha^{-1}}$
$Cov(\hat{\beta}, \hat{\beta}_0) = -(J_\alpha/B_\alpha)^{-1}\left[B_\alpha^{-1}U_\alpha - (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}\xi_\alpha\right]$
$Var(\hat{\beta}) = (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1}$

Remark
The proof of the theorem and of the corollaries below are given in appendix. Let's note that assumption $(A1)$ seems natural and the numerical behaviour is discussed in N'Guessan et al (2001). Theoritical behaviour are displayed for example in Crowder (1984) . Assumption $(A2)$ is designed to support $D_{\alpha^0}^{-1}D_{\hat{\alpha}} \overset{Proba}{\longrightarrow} I_{1+s(r+1)}$ in the proof where alternatives are possible . Matrices $D_{\alpha^0}^{-1}$ and $D_{\hat{\alpha}}$ are also defined in appendix. The behaviour in $(A3)$ is for instance discussed in Crowder (1975) and is used to compute matrix Σ_α . One also notes that $H_\alpha J_\alpha^{-1}H_\alpha^T = R_\alpha$. So using lemma 4, R_α is nonsingular and

$$R_\alpha^{-1} = \Lambda_\alpha^{-1} - \sigma^2(\hat{\beta}_0)\Lambda_\alpha^{-1}\xi_\alpha\xi_\alpha^T\Lambda_\alpha^{-1} \tag{11}$$

Corollary 4.2

$$\sigma^2(\hat{\beta}_0) = ((J_\alpha/B_\alpha) + \parallel \xi_\alpha \parallel^2_{\Lambda_\alpha^{-1}})^{-1} \tag{12}$$

where

$$\parallel \xi_\alpha \parallel^2_{\Lambda_\alpha^{-1}} = \sum_{k=1}^{s} \frac{t_k^2\gamma_k^2}{n_k^2}\frac{\Delta_{\alpha,k}^2}{\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,k}}} \tag{13}$$

with

$$\Delta_{\alpha,k} = z_k^T \Omega_{\alpha,k}^{-1}\mathbf{1}_r \tag{14}$$

and

$$\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\alpha,k}} = \gamma_k^{-1}\left[trace(\Omega_{\alpha,k}^{-1}) + t_k(\frac{\beta_0\gamma_k\Delta_{\alpha,k}}{n_k})^2\right]. \tag{15}$$

Corollary 4.3 *Under the same assumptions than theorem above, the $r \times 1$ covariance vector between the restraints m.l.e. of β_k and β_0 is*

$$Cov(\hat{\beta}_k; \hat{\beta}_0) = \mathcal{S}_{\alpha,k} z_k \tag{16}$$

where $\mathcal{S}_{\alpha,k}$ is the $r \times r$ matrix as follows

$$\mathcal{S}_{\alpha,k} = -\frac{t_k \sigma^2(\hat{\beta}_0)}{n_k} \left[\gamma_k \Omega_{\alpha,k}^{-1} - \frac{1}{\| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}}} \Omega_{\alpha,k}^{-1} (\mathbf{1}_{r,r} + \rho_{\alpha,k} z_k \mathbf{1}_r^T) \Omega_{\alpha,k}^{-1} \right] \tag{17}$$

with

$$\rho_{\alpha,k} = \frac{t_k \Delta_{\alpha,k} \beta_0^2 \gamma_k^2}{n_k^2} \tag{18}$$

for $k = 1, 2, \ldots, s$ and $\mathbf{1}_{r,r} = \mathbf{1}_r \mathbf{1}_r^T$ is the $r \times r$ matrix with unity.

Corollary 4.4 *Under the same assumptions than theorem above, the $r \times r$ between covariance matrix of $\hat{\beta}_k$ and $\hat{\beta}_m$ is*

$$Cov(\hat{\beta}_k; \hat{\beta}_m) = J_{k,m} - \sum_{j=1}^{s} \left[(\sum_{i=!}^{s} A_{i,j} J_{k,i}) \mathbf{1}_{r,r} \right] J_{j,m} \tag{19}$$

k,m = 1,2,...,s. Where $J_{k,k}$ (resp. $J_{k,m}$ for $m \neq k$) is the Schur complements of (J_α / B_α) in the $(1 + r) \times (1 + r)$ matrix

$$\begin{bmatrix} (J_\alpha / B_\alpha) & U_{\alpha,k}^T B_{\alpha,k}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & B_{\alpha,k}^{-1} \end{bmatrix}, \quad (resp. \begin{bmatrix} (J_\alpha / B_\alpha) & U_{\alpha,m}^T B_{\alpha,m}^{-1} \\ B_{\alpha,k}^{-1} U_{\alpha,k} & 0 \end{bmatrix}),$$

and

$$A_{k,m} = \frac{\delta_{k,m}}{\| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}}} - \frac{\sigma^2(\hat{\beta}_0) t_k t_m \gamma_k \gamma_m \Delta_{\alpha,k} \Delta_{\alpha,m}}{n_k n_m \| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}} \| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,m}}}. \tag{20}$$

If $k = m$ then we obtain the within $r \times r$ covariance of $\hat{\beta}_k$ as follows

$$Cov(\hat{\beta}_k; \hat{\beta}_k) = J_{k,k} - \sum_{j=1}^{s} \left[(\sum_{i=1}^{s} A_{i,j} J_{k,i}) \mathbf{1}_{r,r} \right] J_{j,k} \tag{21}$$

k = 1,2,...,s.

5. Examples

5.1. *example 1: $s = 1$ and $r = 2$*

We suppose here that the number of experimental site is equal to one and there are two accident types with n_k the total number of accident of the site and $z_k = (z_{1k}, z_{2k})^T$ the data from the control area. The parameter is defined by $\alpha = (\beta_0, \beta_k)^T \in \mathbb{R}^3$ and $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{jk} > 0$, and $\beta_{1k} + \beta_{2k} = 1$. So the dimension of the parameter space is three and we have one constraint. Following asymptotic results (Crowder 1984 or Neuenschwander & Flury, 1997) and using model (1) and constraints

(2), and after some labouring derivations of the log-likelihood (N'Guessan et al. 2001), we show that the 4×4 constrained Fisher's information matrix is :

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} v_{1k} & \frac{\gamma_k}{\beta_0} v_{2k} & 0 \\ \frac{\gamma_k}{\beta_0} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 1 \\ \frac{\gamma_k}{\beta_0} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

where $\tau_\alpha = \frac{\gamma_k^2 < z_k, \beta_k >}{\beta_0 n_k}$, $v_{mk} = \frac{\gamma_k \beta_0 z_{mk}}{n_k}$, $\gamma_k = \frac{n_k}{1+\beta_0 < z_k, \beta_k >}$, and

$$\begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{21}^{(k)} & b_{22}^{(k)} \end{pmatrix} = \gamma_k \begin{pmatrix} \frac{1+\beta_0 z_{1k}}{\beta_{1k}} - \frac{\beta_0^2 z_{1k}^2 \gamma_k^2}{n_k^2} & -\frac{\beta_0^2 z_{1k}^2 z_{2k}^2 \gamma_k^2}{n_k^2} \\ -\frac{\beta_0^2 z_{2k}^2 z_{1k}^2 \gamma_k^2}{n_k^2} & \frac{1+\beta_0 z_{2k}}{\beta_{2k}} - \frac{\beta_0^2 z_{2k}^2 \gamma_k^2}{n_k^2} \end{pmatrix}.$$

In order to get at the asymptotic covariance matrix of the constrained m.l.e $\tilde{\alpha}$, one must obtain Γ_α^{-1} and take the leading 3×3 submatrix. The method proposed here allows everybody to avoid doing this inversion by using theorem 4.1 as follows;
setting

$$J_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} V_{\alpha,k}^T \\ \frac{\gamma_k}{\beta_0} V_{\alpha,k} & B_{\alpha,k} \end{pmatrix}; \quad \Omega_{\alpha,k} = \begin{pmatrix} \frac{1+\beta_0 z_{1k}}{\beta_{1k}} & 0 \\ 0 & \frac{1+\beta_0 z_{2k}}{\beta_{2k}} \end{pmatrix}$$

and $B_{\alpha,k} = \gamma_k \left[\Omega_{\alpha,k} - V_{\alpha,k} V_{\alpha,k}^T \right]$ a 2×2 matrix and

$$V_{\alpha,k} = \frac{\beta_0 \gamma_k}{n_k} (z_{1k}, z_{2k})^T ; \quad H_2 = (1,1),$$

and using corollary 4.2, the estimated standard error of $\hat{\beta}_0$ is

$$\hat{SE}(\hat{\beta}_0) = \left[(J_{\hat{\alpha}}/B_{\hat{\alpha}}) + \| \xi_{\hat{\alpha}} \|^2_{\Lambda_{\hat{\alpha}}^{-1}} \right]^{-1/2}$$

with

$$(J_{\hat{\alpha}}/B_{\hat{\alpha}}) = \frac{\hat{\gamma}_k^2}{n_k} \left(\frac{< z_k, \hat{\beta}_k >}{\hat{\beta}_0} - \frac{\hat{\gamma}_k \hat{t}_k}{n_k} \| z_k \|^2_{\Omega_{\hat{\alpha},k}^{-1}} \right); \quad \| \xi_{\hat{\alpha}} \|^2_{\Lambda_{\hat{\alpha}}^{-1}} = \frac{\hat{\gamma}_k^2 \hat{t}_k^2 \Delta_{\hat{\alpha},k}^2}{n_k^2 \| \mathbf{1}_r \|^2_{B_{\hat{\alpha},k}^{-1}}}$$

and $\Delta_{\hat{\alpha},k} = \frac{\hat{\beta}_{1k} z_{1k}}{1+\hat{\beta}_0 z_{1k}} + \frac{\hat{\beta}_{2k} z_{2k}}{1+\hat{\beta}_0 z_{2k}}$. Now using the corollary 4.3, one obtains the covariances between the subvector parameter $\hat{\beta}_k$ and the first component $\hat{\beta}_0$ as follows (appendix B)

$$cov(\hat{\beta}_{mk}, \hat{\beta}_0) = \frac{t_k \sigma^2(\hat{\beta}_0)}{n_k} \left(-\frac{\gamma_k \beta_{mk} z_{mk}}{1+\beta_0 z_{mk}} + \frac{(1+\rho_{\alpha,k} z_{mk}) \beta_{mk} \Delta_{\alpha,k}}{(1+\beta_0 z_{mk}) \| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}}} \right) \tag{22}$$

for $m = 1, 2$. Now in view of the proof of corollary 4.4 (*see* appendix B for details) we obtain

$$Var(\hat{\beta}_k) = J_{k,k} - A_{k,k} J_{k,k} \mathbf{1}_{r,r} J_{k,k}$$

a 2×2 matrix and thereafter we deduce

$$\begin{aligned} \sigma^2(\hat{\beta}_{jk}) &= \gamma_k^{-1} \left[\omega_{jk} + \omega_{jk}^2 (\theta_k z_{jk}^2 - A_{k,k} \gamma_k^{-1} X_{jk}^2) \right]; j = 1,2 \\ cov(\hat{\beta}_{1k}; \hat{\beta}_{2k}) &= \gamma_k^{-1} \omega_{1k} \omega_{2k} \left[\theta_k z_{1k} z_{2k} - A_{k,k} \gamma_k^{-1} X_{1k} X_{2k} \right] \end{aligned} \tag{23}$$

ω_{jk} is the diagonal elements of $\Omega_{\alpha,k}^{-1}$ and

$$\theta_k = \frac{t_k \gamma_k^2}{n_k^2} \left(\beta_0^2 - (J_\alpha / B_\alpha)^{-1} t_k \gamma_k\right); \quad X_{jk} = 1 + \theta_k \Delta_{\alpha,k} z_{jk}.$$

5.2. *example 2: $s = 2$ and $r = 2$*

We have two experimental sites and two accident types with n_k, n_m the total number of accident in each site and $z_k = (z_{1k}, z_{2k})^T$, $z_m = (z_{1m}, z_{2m})^T$ the data from each associated control area. The parameter and the constraints are defined by $\alpha = (\beta_0, \beta_k, \beta_m)^T \in \mathbb{R}^5$ where $\beta_k = (\beta_{1k}, \beta_{2k})^T \in \mathbb{R}^2$, $\beta_m = (\beta_{1m}, \beta_{2m})^T \in \mathbb{R}^2$ with $\beta_0 > 0$, $\beta_{ij} > 0$, $\beta_{1k} + \beta_{2k} = 1$ and $\beta_{1m} + \beta_{2m} = 1$. So the dimension of the parameter space is five and we have two constraints. One notes that the additional experimental site to example 1 has extended the dimension of the parameter space by two and the constrained information matrix is a 7×7 matrix given by

$$\Gamma_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} v_{1k} & \frac{\gamma_k}{\beta_0} v_{2k} & \frac{\gamma_m}{\beta_0} v_{1m} & \frac{\gamma_m}{\beta_0} v_{2m} & 0\ 0 \\ \frac{\gamma_k}{\beta_0} v_{1k} & b_{11}^{(k)} & b_{12}^{(k)} & 0 & 0 & 1\ 0 \\ \frac{\gamma_k}{\beta_0} v_{2k} & b_{21}^{(k)} & b_{22}^{(k)} & 0 & 0 & 1\ 0 \\ \frac{\gamma_m}{\beta_0} v_{1m} & 0 & 0 & b_{11}^{(m)} & b_{12}^{(m)} & 0\ 1 \\ \frac{\gamma_m}{\beta_0} v_{1m} & 0 & 0 & b_{21}^{(m)} & b_{22}^{(m)} & 0\ 1 \\ 0 & 1 & 1 & 0 & 0 & 0\ 0 \\ 0 & 0 & 0 & 1 & 1 & 0\ 0 \end{pmatrix}$$

where $\tau_\alpha = \frac{\gamma_k^2 < z_k, \beta_k >}{\beta_0 n_k} + \frac{\gamma_m^2 < z_m, \beta_m >}{\beta_0 n_m}$ and the other terms are defined in the same way as in the example 1. It is the same for vectors $V_{\alpha,k}$, $V_{\alpha,m}$ and matrices $\Omega_{\alpha,k}^{-1}$, and $\Omega_{\alpha,m}^{-1}$. Setting

$$J_\alpha = \begin{pmatrix} \tau_\alpha & \frac{\gamma_k}{\beta_0} V_{\alpha,k}^T & \frac{\gamma_m}{\beta_0} V_{\alpha,m}^T \\ \frac{\gamma_k}{\beta_0} V_{\alpha,k} & B_{\alpha,k} & \bigcirc \\ \frac{\gamma_m}{\beta_0} V_{\alpha,m} & \bigcirc & B_{\alpha,m} \end{pmatrix}, \quad B_\alpha = \begin{pmatrix} B_{\alpha,k} & \bigcirc \\ \bigcirc & B_{\alpha,m} \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1\ 1\ 0\ 0 \\ 0\ 0\ 1\ 1 \end{pmatrix}$$

and in view of corollary 4.2, the estimated standard error of the first component is

$$\hat{SE}(\hat{\beta}_0) = \left[(J_{\hat{\alpha}} / B_{\hat{\alpha}}) + \| \xi_{\hat{\alpha}} \|^2_{A_{\hat{\alpha}}^{-1}}\right]^{-1/2}$$

with

$$(J_{\hat{\alpha}} / B_{\hat{\alpha}}) = \sum_{i \in \{k,m\}} \frac{\hat{\gamma}_i^2}{n_i} \left(\frac{< z_i, \hat{\beta}_i >}{\hat{\beta}_0} - \frac{\hat{\gamma}_i \hat{t}_i}{n_i} \| z_i \|^2_{\Omega_{\hat{\alpha},i}^{-1}} \right); \quad \| \xi_{\hat{\alpha}} \|^2_{A_{\hat{\alpha}}^{-1}} = \sum_{i \in \{k,m\}} \frac{\hat{\gamma}_i^2 \hat{t}_i^2 \Delta_{\hat{\alpha},i}^2}{n_i^2 \| \mathbf{1}_r \|^2_{B_{\alpha,i}^{-1}}}.$$

The covariances between $\hat{\beta}_0$ and the two other estimators $\hat{\beta}_k$ and $\hat{\beta}_m$ are obtained in the same way as in example 1. We have here (*see* appendix B) one between covariances matrix $Cov(\hat{\beta}_k; \hat{\beta}_m)$ and two within covariances matrices $Var(\hat{\beta}_k)$ and $Var(\hat{\beta}_m)$. So the standard errors of $\hat{\beta}_{jk}$ and $\hat{\beta}_{jm}$ are obtained with the square root of the diagonal elements of each within covariances matrix. We don't need any numerical

inversion programm to succed in doing it, only matrices manipulations (product and addition) are called for. For example (*see* for furthers technical details in appendix B) the 2×2 within matrix $Var(\hat{\beta}_k)$ is obtained as folows

$$Var(\hat{\beta}_k) = \gamma_k^{-1}\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1}\left[\phi_{k,m} z_k z_k^T + a_{k,m}(z_k \mathbf{1}_r^T + \mathbf{1}_r z_k^T) - A_{k,k}\gamma_k^{-2} X_k X_k^T\right]\Omega_{\alpha,k}^{-1}$$

where

$$\phi_{k,m} = \gamma_k^{-1}\theta_k + 2A_{k,m}v_{k,m}\gamma_k^{-1}\theta_k - A_{m,m}v_{k,m}^2\Delta_{\alpha,k}^{-2}, \quad a_{k,m} = A_{k,m}v_{k,m}\gamma_k^{-1}\Delta_{\alpha,k}^{-1}.$$

and

$$v_{k,m} = (J_\alpha/B_\alpha)^{-1}\frac{t_k\gamma_k\Delta_{\alpha,k}}{n_k}\frac{t_m\gamma_m\Delta_{\alpha,m}}{n_m}.$$

6. Concluding remarks

We can use any inversion method (Guttman 1946, ouellette 1981 pp. 255 - 266) or any computer programm (Press et al 1992, N'Guessan 1993) to calculate numerically and directly the asymptotic covariances matrix W_α (*see* appendix) by inverting the matrix

$$\Gamma_\alpha = \begin{pmatrix} J_\alpha & H_\alpha^T \\ H_\alpha & 0 \end{pmatrix}$$

and thereafter standard errors are obtained by taking the square root of the diagonal elements of the covariance matrix equal to the leading $(1+sr) \times (1+sr)$ submatrix in Γ_α^{-1} i.e. W_α. However one recalls that the dimension of the parameter space and the number of constrained related to the model above are very closed to s and r . So the size of matrix Γ_α increases with them i.e. with the dimension of accident data tables which are to be combined. Of course this latter relation may seriously affect the numerical values of Γ_α^{-1} when s and r are large, and sometimes leads to instability in the coefficients of Γ_α^{-1} and therefore in those of W_α. The main strength of our approach is its applicability to any dimension of matrix Γ_α without being obliged to invert it. It saves a considerable time and precision in calculations and is accessible to everyone. Theorem 4.1 provides a powerful tool to compute asymptotic variances and covariances based on relatively straightforward calculations and a reduction of dimensionality. Once data z_k, n_k have been given and scalars γ_k, τ_α and vectors $V_{\alpha,k}$, and matrices $\Omega_{\alpha,k}$ have been found, theorem 4.1 gives a general formula for the asymptotic covariance matrix W_α. Moreover, if $\hat{W}_\alpha$ denotes the matrix W_α evaluated at the restraint maximum likelihood estimate $\hat{\alpha}$, then estimated variances and covariances of the parameter estimates can be easily computed, estimated standard errors are obtained by taking the square root of the diagonal elements of the same matrix, and asymptotic normal theory confidence regions can be constructed as well. For example, the standard error of β_0 the element of interest of $\hat{\alpha}$ is obtained as follows

$$\mathrm{SE}(\hat{\beta}_0) = \left[(J_{\hat{\alpha}}/B_{\hat{\alpha}}) + \| \xi_{\hat{\alpha}} \|^2_{\Lambda_{\hat{\alpha}}^{-1}}\right]^{-1/2}$$

and without matrix inversion. We can also get asymptotic correlation coefficients by capitalizing on the elements apart from the diagonal. Furthers numerical developments will make capital out of those results and remarks at subsequent papers in order to bring forward asymptotic confidence intervals and correlation coefficients for all components of the restraint m.l.e. $\hat{\alpha}$.

Appendix A : Proof of the main result

Proof. of the theorem 4.1
Application of the mean value theorem at α^0 to (10) gives

$$(\frac{\partial L}{\partial \alpha})_{\alpha^0} + (\frac{\partial^2 L}{\partial \alpha \partial \alpha})_{\alpha}(\hat{\alpha} - \alpha^0) + H_{\alpha}^T \hat{\lambda} = 0$$

$$H_{\alpha}(\hat{\alpha} - \alpha^0) = 0. \tag{24}$$

Now combining lemma 4 and the constraints (2), then likelihood equations can be written as

$$(\frac{\partial L}{\partial \alpha})_{\alpha^0} + J_{\alpha}(M_{\alpha} - I_{1+sr})(\hat{\alpha} - \alpha^0) + H_{\alpha}^T \hat{\lambda} = 0$$

$$H_{\alpha}(\hat{\alpha} - \alpha^0) = 0. \tag{25}$$

A matrix version of the latter is as follows

$$D_{\alpha} \begin{pmatrix} \hat{\alpha} - \alpha^0 \\ \eta_{\hat{\alpha}} \end{pmatrix} = \begin{pmatrix} J_{\alpha}^{-1}(\frac{\partial L}{\partial \alpha})_{\alpha^0} \\ 0 \end{pmatrix} \tag{26}$$

where

$$D_{\alpha} = \begin{pmatrix} I_{1+sr} - M_{\alpha} & -\Phi_{\alpha} \\ H_{\alpha} & 0_s \end{pmatrix}; \quad \eta_{\hat{\alpha}} = H_2(J_{\alpha}/\tau_{\alpha})^{-1} H_2^T \hat{\lambda}.$$

The matrix J_{α}^{-1} and D_{α} are evaluated at any point of the line segment joining α^0 and $\hat{\alpha}$. Under $(A1)$ and $(A2)$ and using a version of a theorem of Crowder (1984, section 2), then the latter system of equations can be written as

$$\begin{pmatrix} \hat{\alpha} - \alpha^0 \\ \eta_{\hat{\alpha}} \end{pmatrix} \underset{\sim}{\mathcal{D}_{1+sr.}} D_{\alpha^0}^{-1} \begin{pmatrix} J_{\alpha}^{-1}(\frac{\partial L}{\partial \alpha})_{\alpha^0} \\ 0 \end{pmatrix} \tag{27}$$

where

$$D_{\alpha^0} = \begin{pmatrix} I_{1+sr} & -\Phi_{\alpha^0} \\ H_{\alpha^0} & 0_s \end{pmatrix}; \quad D_{\alpha^0}^{-1} = \begin{pmatrix} I_{1+sr} - \Phi_{\alpha^0} H_{\alpha^0} & \Phi_{\alpha^0} \\ -H_{\alpha^0} & I_s \end{pmatrix}.$$

By $(A3)$ the asymptotic distribution of the right-hand side of (27) is $\mathcal{N}(0, \Sigma_{\alpha^0})$ where

$$\Sigma_{\alpha^0} = D_{\alpha^0}^{-1} \begin{pmatrix} J_{\alpha^0}^{-1} & 0 \\ 0 & 0 \end{pmatrix} (D_{\alpha^0}^{-1})^T.$$

After some matrix manipulation, we obtain the formal expression of Σ_{α} as follows

$$\Sigma_{\alpha} = \begin{pmatrix} W_{\alpha} & 0 \\ 0 & R_{\alpha} \end{pmatrix}; \quad W_{\alpha} = J_{\alpha}^{-1} - J_{\alpha}^{-1} H_{\alpha}^T R_{\alpha}^{-1} H_{\alpha} J_{\alpha}^{-1}.$$

Now using part $i)$ of lemma 4 and remindering that $H_1 = O_s$ we get:

$$J_{\alpha}^{-1} H_{\alpha}^T = \begin{pmatrix} -(J_{\alpha}/B_{\alpha})^{-1} \xi_{\alpha}^T \\ (J_{\alpha}/\tau_{\alpha})^{-1} H_2^T \end{pmatrix}$$

a $(1+sr)\times s$ matrix and

$$H_\alpha^T J_\alpha^{-1} = \left(-(J_\alpha/B_\alpha)^{-1}\xi_\alpha, H_2(J_\alpha/\tau_\alpha)^{-1}\right)$$

a $s\times(1+sr)$ matrix. One deduces the formal compoments of W_α as follows

$$\begin{aligned} W_\alpha = &\begin{bmatrix} (J_\alpha/B_\alpha)^{-1} & -(J_\alpha/B_\alpha)^{-1}U_\alpha^T B_\alpha^{-1} \\ -B_\alpha^{-1}U_\alpha(J_\alpha/B_\alpha)^{-1} & (J_\alpha/\tau_\alpha)^{-1} \end{bmatrix} - \begin{bmatrix} -(J_\alpha/B_\alpha)^{-1}\xi_\alpha^T \\ (J_\alpha/\tau_\alpha)^{-1}H_2^T \end{bmatrix} \times \\ &R_\alpha^{-1}\left[-(J_\alpha/B_\alpha)^{-1}\xi_\alpha, H_2(J_\alpha/\tau_\alpha)^{-1}\right]. \end{aligned} \tag{28}$$

After some manipulations, the formal expressions for the variances covariances of $\hat{\alpha}$ are

$$\begin{aligned} W_\alpha^{(11)} &= (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-1}\xi_\alpha^T R_\alpha^{-1}\xi_\alpha(J_\alpha/B_\alpha)^{-1} \\ W_\alpha^{(12)} &= (J_\alpha/B_\alpha)^{-1}U_\alpha^T B_\alpha^{-1} + (J_\alpha/B_\alpha)^{-1}\xi_\alpha^T R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1} \\ W_\alpha^{(21)} &= -B_\alpha^{-1}U_\alpha(J_\alpha/B_\alpha)^{-1} + (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}\xi_\alpha(J_\alpha/B_\alpha)^{-1} \\ W_\alpha^{(22)} &= (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1}. \end{aligned} \tag{29}$$

■

Proof. of Corollary 4.2
Using lemma 4 part *ii*), one obtains

$$\begin{aligned} R_\alpha^{-1}\xi_\alpha &= \left(\Lambda_\alpha^{-1} - ((J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}})^{-1}\Lambda_\alpha^{-1}\xi_\alpha\xi_\alpha^T\Lambda_\alpha^{-1}\right)\xi_\alpha \\ &= \left(1 - ((J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}})^{-1} \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}\right)\Lambda_\alpha^{-1}\xi_\alpha \\ &= \frac{(J_\alpha/B_\alpha)}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}}\Lambda_\alpha^{-1}\xi_\alpha \end{aligned} \tag{30}$$

then

$$\begin{aligned} \| \xi_\alpha \|^2_{R_\alpha^{-1}} &= \xi_\alpha^T R_\alpha^{-1}\xi_\alpha \\ &= \frac{(J_\alpha/B_\alpha)}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}} \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}. \end{aligned} \tag{31}$$

We deduce

$$\begin{aligned} \sigma^2(\hat{\beta}_0) &= (J_\alpha/B_\alpha)^{-1} - (J_\alpha/B_\alpha)^{-2} \| \xi_\alpha \|^2_{R_\alpha^{-1}} \\ &= \frac{1}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}}. \end{aligned} \tag{32}$$

The explicit expressions of $\| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}$ and $\| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}}$ provide from a formal application of lemma 2 as follows

$$
\begin{aligned}
\| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}} &= \xi_\alpha^T \Lambda_\alpha^{-1} \xi_\alpha \\
&= U_\alpha^T B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} H_2 B_\alpha^{-1} U_\alpha \\
&= (U_{\alpha,1}^T B_{\alpha,1}^{-1}, \ldots, U_{\alpha,s}^T B_{\alpha,s}^{-1})\, H_2^T \Lambda_\alpha^{-1} H_2\, (B_{\alpha,1}^{-1} U_{\alpha,1}, \ldots, B_{\alpha,s}^{-1} U_{\alpha,s}) \\
&= (U_{\alpha,1}^T B_{\alpha,1}^{-1}, \ldots, U_{\alpha,s}^T B_{\alpha,s}^{-1}) \begin{bmatrix} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,1}^{-1}}} & \bigcirc & \cdots & \bigcirc \\ \bigcirc & \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,2}^{-1}}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \cdots & \bigcirc & \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,s}^{-1}}} \end{bmatrix} (B_{\alpha,1}^{-1} U_{\alpha,1}, \ldots, B_{\alpha,s}^{-1} U_{\alpha,s}) \\
&= \sum_{k=1}^{s} U_{\alpha,k}^T B_{\alpha,k}^{-1} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}}} B_{\alpha,k}^{-1} U_{\alpha,k} \\
&= \sum_{k=1}^{s} \frac{t_k^2 \gamma_k^2}{n_k^2} z_k^T \Omega_{\alpha,k}^{-1} \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}}} \Omega_{\alpha,k}^{-1} z_k \\
\| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}} &= \sum_{k=1}^{s} \frac{t_k^2 \gamma_k^2}{n_k^2} \frac{\Delta_{\alpha,k}^2}{\| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}}},
\end{aligned}
\tag{33}
$$

and

$$
\begin{aligned}
\| \mathbf{1}_r \|^2_{B_{\alpha,k}^{-1}} &= \mathbf{1}_r^T B_{\alpha,k}^{-1} \mathbf{1}_r \\
&= \mathbf{1}_r^T \gamma_k^{-1} (\Omega_{\alpha,k}^{-1} + t_k \Omega_{\alpha,k}^{-1} V_{\alpha,k} V_{\alpha,k}^T \Omega_{\alpha,k}^{-1}) \mathbf{1}_r \\
&= \gamma_k^{-1} [\mathrm{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k^2}{n_k^2} \mathbf{1}_r^T \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1} \mathbf{1}_r] \\
&= \gamma_k^{-1} [\mathrm{trace}(\Omega_{\alpha,k}^{-1}) + t_k \frac{\beta_0^2 \gamma_k^2 \Delta_{\alpha,k}^2}{n_k^2}]
\end{aligned}
\tag{34}
$$

■ **Proof.** of Corollary 4.3

Notice that

$$
Cov(\hat{\beta}; \hat{\beta}_0) = W_\alpha^{(21)}
$$

so using lemma 3 and 4 we deduce

$$
\begin{aligned}
Cov(\hat{\beta}; \hat{\beta}_0) &= -\frac{1}{(J_\alpha / B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}} [B_\alpha^{-1} U_\alpha - B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} \xi_\alpha] \\
&= -\frac{1}{(J_\alpha / B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}} [I_{sr} - B_\alpha^{-1} H_2^T \Lambda_\alpha^{-1} H_2] B_\alpha^{-1} U_\alpha
\end{aligned}
\tag{35}
$$

with

$$
H_2^T \Lambda_\alpha^{-1} H_2 = \text{Bloc-diag}(\frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,1}^{-1}}}, \ldots, \frac{\mathbf{1}_r \mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B_{\alpha,s}^{-1}}})
$$

a $(sr) \times (sr)$ matrix where each bloc $k (k = 1, 2, \ldots, s)$ is a $r \times r$ matrix given by the expression

$$\frac{\mathbf{1}_r\mathbf{1}_r^T}{\| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}}} = \frac{1}{\| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}}} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}. \tag{36}$$

Hence

$$Cov(\hat{\beta}_k; \hat{\beta}_0) = -\frac{1}{(J_\alpha/B_\alpha) + \| \xi_\alpha \|^2_{\Lambda_\alpha^{-1}}} (I_r - B^{-1}_{\alpha,k} \frac{\mathbf{1}_{r,r}}{\| \mathbf{1}_r \|^2_{B^{-1}_{\alpha,k}}}) B^{-1}_{\alpha,k} U_{\alpha,k} \tag{37}$$

for $k = 1, 2, \dots, s$. Now using part i) of lemma 2 we show that

$$B^{-1}_{\alpha,k} U_{\alpha,k} = \frac{t_k \gamma_k}{n_k} \Omega^{-1}_{\alpha,k} z_k \tag{38}$$

and

$$B^{-1}_{\alpha,k} \mathbf{1}_{r,r} \Omega^{-1}_{\alpha,k} = \gamma_k^{-1} \left[\Omega^{-1}_{\alpha,k} (\mathbf{1}_r \mathbf{1}_r^T + \rho_{\alpha,k} z_k \mathbf{1}_r^T) \Omega^{-1}_{\alpha,k} \right] \tag{39}$$

So after some algebraic manipulations, one obtains the expression of matrix $\mathcal{S}_{\alpha,k}$ and the result of the corollary. ■

Proof. of Corollary 4.4
This proof is an algebriac manipulations of matrices. So we just give the outline. From theorem 4.1 $Var(\hat{\beta})$ is the $(sr) \times (sr)$ matrix

$$Var(\hat{\beta}) = (J_\alpha/\tau_\alpha)^{-1} - (J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1}.$$

Now using lemma 3 and matrix $B_\alpha^{-1} = \text{Bloc-diag}(B^{-1}_{\alpha,1}, \dots, B^{-1}_{\alpha,s})$, one shows that

$$(J_\alpha/\tau_\alpha)^{-1} = \begin{bmatrix} B^{-1}_{\alpha,1} & \bigcirc & \dots & \bigcirc \\ \bigcirc & B^{-1}_{\alpha,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bigcirc \\ \bigcirc & \dots & \bigcirc & B^{-1}_{\alpha,s} \end{bmatrix} - (J_\alpha/B_\alpha)^{-1} \begin{bmatrix} B^{-1}_{\alpha,1} U_{\alpha,1} \\ B^{-1}_{\alpha,2} U_{\alpha,2} \\ \vdots \\ B^{-1}_{\alpha,s} U_{\alpha,s} \end{bmatrix} \left[U^T_{\alpha,1} B^{-1}_{\alpha,1}, U^T_{\alpha,2} B^{-1}_{\alpha,2}, \dots U^T_{\alpha,s} B^{-1}_{\alpha,s} \right]$$

So $(J_\alpha/\tau_\alpha)^{-1} = ((J_{k,m}))$; $k, m = 1, 2, \dots, s$ a $(sr) \times (sr)$ matrix whose entry is also a $r \times r$ matrix such that

$$J_{k,m} = \delta_{k,m} B^{-1}_{\alpha,k} - (J_\alpha/B_\alpha)^{-1} B^{-1}_{\alpha,k} U_{\alpha,k} U^T_{\alpha,m} B^{-1}_{\alpha,m}. \tag{40}$$

In view of (11), and the relations

$$\begin{aligned} \mathbf{1}_{r,r} B^{-1}_{\alpha,k} U_{\alpha,k} U^T_{\alpha,m} B^{-1}_{\alpha,m} \mathbf{1}_{r,r} &= \frac{t_k t_m \Delta_{\alpha,k} \Delta_{\alpha,m}}{n_k n_m} \mathbf{1}_{r,r} \\ \Delta_{\alpha,k} = z_k^T \Omega^{-1}_{\alpha,k} \mathbf{1}_r &= \mathbf{1}_r^T \Omega^{-1}_{\alpha,k} z_k \end{aligned} \tag{41}$$

we deduce

$$H_2^T R_\alpha^{-1} H_2 = A_\alpha \otimes \mathbf{1}_{r,r} \tag{42}$$

where $\otimes$ is the Kronecker product, $A_\alpha = ((A_{k,m}))$ a $s \times s$ matrix whose entry is given by (20). So postmultiplication and premultiplication of $H_2^T R_\alpha^{-1} H_2$ by $(J_\alpha/\tau_\alpha)^{-1}$, we get

$$(J_\alpha/\tau_\alpha)^{-1}H_2^T R_\alpha^{-1}H_2(J_\alpha/\tau_\alpha)^{-1} =$$

$$= \begin{bmatrix} \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{1,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,1} & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{1,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,2} & \dots & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{1,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,s} \\ \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{2,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,1} & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{2,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,2} & \dots & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{2,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,s} \\ \dots & \dots & \dots & \dots \\ \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{s,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,1} & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{s,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,2} & \dots & \sum_{j=1}^{s}[(\sum_{i=1}^{s} J_{s,i}A_{i,j})\mathbf{1}_{r,r}]J_{j,s} \end{bmatrix} \quad (43)$$

So, we obtain the second term of the right hand of (19). The result follows from matrices addition. ■

Appendix B: Some technical details for examples 1 and 2

Example 1

For $s = 1$ and $r > 1$, and in view of lemma 2 part i), the $r \times r$ Schur complement matrix $J_{k,k}$ is

$$J_{k,k} = \gamma_k^{-1}\left[\Omega_{\alpha,k}^{-1} + \theta_k \Omega_{\alpha,k}^{-1} z_k z_k^T \Omega_{\alpha,k}^{-1}\right]. \quad (44)$$

So using the latter expression of $J_{k,k}$ and after some matrices manipulations, one obtains

$$\begin{aligned} Var(\hat{\beta}_k) &= J_{k,k} - A_{k,k}J_{k,k}\mathbf{1}_{r,r}J_{k,k} \\ &= \gamma_k^{-1}\left[\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1}(\theta_k z_k z_k^T - A_{k,k}\gamma_k^{-1}X_k X_k^T)\Omega_{\alpha,k}^{-1}\right] \end{aligned} \quad (45)$$

where $X_k = (X_{1k}, \dots, X_{rk})^T$ a $r \times 1$ vector of elements $X_{jk} = 1 + \theta_k \Delta_{\alpha,k} z_{jk}$. In this way, we deduce

$$\begin{aligned} \sigma^2(\hat{\beta}_{jk}) &= \gamma_k^{-1}\left[\omega_{jk} + \omega_{jk}^2(\theta_k z_{jk}^2 - A_{k,k}\gamma_k^{-1}X_{jk}^2)\right]; j = 1, 2, \dots, r \\ cov(\hat{\beta}_{jk}; \hat{\beta}_{mk}) &= \gamma_k^{-1}\omega_{jk}\omega_{mk}\left[\theta_k z_{jk} z_{mk} - A_{k,k}\gamma_k^{-1}X_{jk}X_{mk}\right]; j \neq m = 1, 2, \dots, r. \end{aligned} \quad (46)$$

A straightforword development of the $r \times r$ matrix $\mathcal{S}_{\alpha,k}$ of corollary 4.3 implies the covariance between $\hat{\beta}_0$ and any other estimator $\hat{\beta}_{mk}$ as follows

$$\begin{aligned} \frac{n_k}{t_k\sigma^2(\hat{\beta}_0)}cov(\hat{\beta}_{mk}, \hat{\beta}_0) = &-\left(\frac{\gamma_k\beta_{mk}}{1+\beta_0 z_{mk}} - \frac{(1+\rho_{\alpha,k}z_{mk})\beta_{mk}^2}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{mk})^2}\right) z_{mk} + \\ &+ \frac{(1+\rho_{\alpha,k}z_{mk})\beta_{mk}}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{mk})}\left(\Delta_{\alpha,k} - \frac{\beta_{mk}z_{mk}}{1+\beta_0 z_{mk}}\right) \end{aligned} \quad (47)$$

for $m = 1, 2, \dots, r$.

Now for $r = 2$,

$$\mathcal{S}_{\alpha,k} = -\frac{t_k\sigma^2(\hat{\beta}_0)}{n_k}\begin{pmatrix} \frac{\gamma_k\beta_{1k}}{1+\beta_0 z_{1k}} - \frac{(1+\rho_{\alpha,k}z_{1k})\beta_{1k}^2}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{1k})^2} & -\frac{(1+\rho_{\alpha,k}z_{1k})\beta_{1k}\beta_{2k}}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{1k})(1+\beta_0 z_{2k})} \\ -\frac{(1+\rho_{\alpha,k}z_{2k})\beta_{2k}\beta_{1k}}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{2k})(1+\beta_0 z_{1k})} & \frac{\gamma_k\beta_{2k}}{1+\beta_0 z_{2k}} - \frac{(1+\rho_{\alpha,k}z_{2k})\beta_{2k}^2}{\| \mathbf{1}_r \|^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{2k})^2} \end{pmatrix}$$

we get more details for the covariances between the 2×1 vector $\hat{\beta}_k$ and β_0 of example 1 as follows

$$\frac{n_k}{t_k\sigma^2(\hat{\beta}_0)}cov(\hat{\beta}_{1k},\hat{\beta}_0) = -\left(\frac{\gamma_k\beta_{1k}}{1+\beta_0 z_{1k}} - \frac{(1+\rho_{\alpha,k}z_{1k})\beta_{1k}^2}{\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{1k})2}\right) z_{1k} + \\ + \left(\frac{(1+\rho_{\alpha,k}z_{1k})\beta_{1k}\beta_{2k}}{\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{1k})(1+\beta_0 z_{2k})}\right) z_{2k} \tag{48}$$

and

$$\frac{n_k}{t_k\sigma^2(\hat{\beta}_0)}cov(\hat{\beta}_{2k},\hat{\beta}_0) = \left(\frac{(1+\rho_{\alpha,k}z_{2k})\beta_{1k}\beta_{2k}}{\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{1k})(1+\beta_0 z_{2k})}\right) z_{1k} - \\ - \left(\frac{\gamma_k\beta_{2k}}{1+\beta_0 z_{2k}} - \frac{(1+\rho_{\alpha,k}z_{2k})\beta_{2k}^2}{\parallel \mathbf{1}_r \parallel^2_{B^{-1}_{\hat{\alpha},k}}(1+\beta_0 z_{2k})^2}\right) z_{2k}. \tag{49}$$

Example 2

For $s = 2$ (two experimental sites noted k and m) and $r > 1$, A_α (resp. $H_2^T R_\alpha^{-1} H_2$) is a 2×2 (resp. $2r \times 2r$) matrix given by

$$A_\alpha = \begin{bmatrix} A_{k,k} & A_{k,m} \\ A_{k,m} & A_{m,m} \end{bmatrix}, \quad (\text{resp.} \quad H_2^T R_\alpha^{-1} H_2 = \begin{bmatrix} A_{k,k}\mathbf{1}_{r,r} & A_{k,m}\mathbf{1}_{r,r} \\ A_{k,m}\mathbf{1}_{r,r} & A_{m,m}\mathbf{1}_{r,r} \end{bmatrix}).$$

Therefore we have

$$(J_\alpha/\tau_\alpha)^{-1} = \begin{bmatrix} J_{k,k} & J_{k,m} \\ J_{m,k} & J_{m,m} \end{bmatrix} \tag{50}$$

a $2r \times 2r$ matrix. So,

$$(J_\alpha/\tau_\alpha)^{-1} H_2^T R_\alpha^{-1} H_2 (J_\alpha/\tau_\alpha)^{-1} = \begin{bmatrix} \sum\limits_{(j,i)\in\{k,m\}} J_{k,i}A_{i,j}\mathbf{1}_{r,r}J_{j,k} & \sum\limits_{(j,i)\in\{k,m\}} J_{k,i}A_{i,j}\mathbf{1}_{r,r}J_{j,m} \\ \sum\limits_{(j,i)\in\{k,m\}} J_{m,i}A_{i,j}\mathbf{1}_{r,r}J_{j,k} & \sum\limits_{(j,i)\in\{k,m\}} J_{m,i}A_{i,j}\mathbf{1}_{r,r}J_{j,m} \end{bmatrix}$$

and we deduce two $r \times r$ within covariances matrix and one $r \times r$ between covariance matrix as follows

$$Var(\hat{\beta}_k) = J_{k,k} - \sum_{(j,i)\in\{k,m\}} J_{k,i}A_{i,j}\mathbf{1}_{r,r}J_{j,k} \tag{51}$$

$$Var(\hat{\beta}_m) = J_{m,m} - \sum_{(j,i)\in\{k,m\}} J_{m,i}A_{i,j}\mathbf{1}_{r,r}J_{j,m} \tag{52}$$

and

$$Cov(\hat{\beta}_k;\hat{\beta}_m) = J_{k,m} - \sum_{(j,i)\in\{k,m\}} J_{k,i}A_{i,j}\mathbf{1}_{r,r}J_{j,m}, \quad k \neq m. \tag{53}$$

For example

$$\begin{aligned} Var(\hat{\beta}_k) &= J_{k,k} - \sum_{(j,i)\in\{k,m\}} J_{k,i}A_{i,j}\mathbf{1}_{r,r}J_{j,k} \\ &= J_{k,k} - A_{k,k}J_{k,k}\mathbf{1}_{r,r}J_{k,k} - \\ &- A_{k,m}J_{k,m}\mathbf{1}_{r,r}J_{k,k} - A_{k,m}J_{k,k}\mathbf{1}_{r,r}J_{m,k} - A_{m,m}J_{k,m}\mathbf{1}_{r,r}J_{m,k}. \end{aligned} \tag{54}$$

Setting

$$v_{k,m} = (J_\alpha/B_\alpha)^{-1} \frac{t_k \gamma_k \Delta_{\alpha,k}}{n_k} \frac{t_m \gamma_m \Delta_{\alpha,m}}{n_m}$$

then

$$\begin{aligned} J_{k,m}\mathbf{1}_{r,r}J_{k,k} &= -v_{k,m}\gamma_k^{-1}\Delta_{\alpha,k}^{-1}\Omega_{\alpha,k}^{-1}(z_k\mathbf{1}_r^T + \theta_k\Delta_{\alpha,k}z_kz_k^T)\Omega_{\alpha,k}^{-1} \\ J_{k,k}\mathbf{1}_{r,r}J_{m,k} &= -v_{k,m}\gamma_k^{-1}\Delta_{\alpha,k}^{-1}\Omega_{\alpha,k}^{-1}(\mathbf{1}_r z_k^T + \theta_k\Delta_{\alpha,k}z_kz_k^T)\Omega_{\alpha,k}^{-1} \\ J_{k,m}\mathbf{1}_{r,r}J_{m,k} &= v_{k,m}^2\Delta_{\alpha,k}^{-2}\Omega_{\alpha,k}^{-1}z_kz_k^T\Omega_{\alpha,k}^{-1}. \end{aligned} \tag{55}$$

So we deduce

$$\begin{aligned} Var(\hat{\beta}_k) = \gamma_k^{-1}&\left[\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1}(\theta_k z_k z_k^T - A_{k,k}\gamma_k^{-1}X_kX_k^T)\Omega_{\alpha,k}^{-1}\right] + \\ &+ A_{k,m}v_{k,m}\gamma_k^{-1}\Delta_{\alpha,k}^{-1}\Omega_{\alpha,k}^{-1}\left[z_k\mathbf{1}_r^T + \mathbf{1}_r z_k^T + 2\theta_k\Delta_{\alpha,k}z_kz_k^T\right]\Omega_{\alpha,k}^{-1} - \\ &- A_{m,m}v_{k,m}^2\Delta_{\alpha,k}^{-2}\Omega_{\alpha,k}^{-1}z_kz_k^T\Omega_{\alpha,k}^{-1} \end{aligned} \tag{56}$$

where

$$z_k\mathbf{1}_r^T + \mathbf{1}_r z_k^T = \begin{bmatrix} 2z_{1k} & z_{1k}+z_{2k} & \dots & z_{1k}+z_{rk} \\ z_{2k}+z_{1k} & 2z_{2k} & \dots & z_{2k}+z_{rk} \\ \vdots & \ddots & \ddots & \vdots \\ z_{rk}+z_{1k} & \dots & & 2z_{rk} \end{bmatrix}; \quad z_kz_k^T = \begin{bmatrix} z_{1k}^2 & z_{1k}z_{2k} & \dots & z_{1k}z_{rk} \\ z_{2k}z_{1k} & z_{2k}^2 & \dots & z_{2k}z_{rk} \\ \vdots & \ddots & \ddots & \vdots \\ z_{rk}z_{1k} & \dots & & z_{rk}^2 \end{bmatrix}$$

and the $r \times r$ matrix $X_kX_k^T$ is given as follows

$$\begin{bmatrix} (1+\theta_k\Delta_{\alpha,k}z_{1k})^2 & (1+\theta_k\Delta_{\alpha,k}z_{1k})(1+\theta_k\Delta_{\alpha,k}z_{2k}) & \dots & (1+\theta_k\Delta_{\alpha,k}z_{1k})(1+\theta_k\Delta_{\alpha,k}z_{rk}) \\ (1+\theta_k\Delta_{\alpha,k}z_{1k})(1+\theta_k\Delta_{\alpha,k}z_{2k}) & (1+\theta_k\Delta_{\alpha,k}z_{2k})^2 & \dots & (1+\theta_k\Delta_{\alpha,k}z_{2k})(1+\theta_k\Delta_{\alpha,k}z_{rk}) \\ \vdots & \ddots & \ddots & \vdots \\ (1+\theta_k\Delta_{\alpha,k}z_{1k})(1+\theta_k\Delta_{\alpha,k}z_{rk}) & \dots & & (1+\theta_k\Delta_{\alpha,k}z_{rk})^2 \end{bmatrix}.$$

So

$$Var(\hat{\beta}_k) = \gamma_k^{-1}\Omega_{\alpha,k}^{-1} + \Omega_{\alpha,k}^{-1}\left[\phi_{k,m}z_kz_k^T + a_{k,m}(z_k\mathbf{1}_r^T + \mathbf{1}_r z_k^T) - A_{k,k}\gamma_k^{-2}X_kX_k^T\right]\Omega_{\alpha,k}^{-1}$$

where

$$\phi_{k,m} = \gamma_k^{-1}\theta_k + 2A_{k,m}v_{k,m}\gamma_k^{-1}\theta_k - A_{m,m}v_{k,m}^2\Delta_{\alpha,k}^{-2}, \quad a_{k,m} = A_{k,m}v_{k,m}\gamma_k^{-1}\Delta_{\alpha,k}^{-1}.$$

acknowledgement

I am especially grateful to Prof. Claude Langrand for his helpful comments and suggestions, which have led to an improvment in this article which follows from my Ph.D thesis in road safety measure analysis. I wish also thank Professor Francois Bellavance of HEC Montreal for helpful discussions. Part of this work was carried out while the author was an invited researcher of CRT (Centre for Research on Transportation) of Université de Montréal, whose partial support is gratefully acknowledged.

References

[1] Abdel-Aty M.A., Radwan A.E (2000), *Modeling traffic accident occurrence and involvement*, Accid. Anal. and Prev. 32, 633-642.

[2] Aitchison J., Silvey S.D. (1958) *Maximum Likelihood Estimation of parameters subject to restraints* Annals of mathematical statistics, 29,813-829.

[3] Barry S., Ginpil S, O'Neill, T.J (1999), *The effectiveness of air bags*, Accid. Anal. and Prev. 31, 781-787.

[4] Beenstock M., Gafni D., Goldin E (2001), *The effect of traffic policing on road safety in Israel*, Accid. Anal. and Prev. 33, 73-80.

[5] Crowder M. (1975) *Maximum likelihood estimation for dependent observations* J. R. Statist. Soc., 38, 45-53

[6] Crowder M. (1984) *On the constrained maximum likelihood estimation with non i.i.d. observations* Ann. Inst. Statist. Math. 36, A, 239-249.

[7] Danielson S. (1988), *Estimation of the effects of countermeasures on differents types of accidents in the presence of regression effects*, Accid. Anal. and Prev. 20, $n°4$, 289–298.

[8] Danielson S. (1986), *A comparison of two methods for estimating of the effect of a countermeasure in the presence of regression effects*, Accid. Anal. and Prev. 18, $n°1$, 13–23.

[9] Frees E.W (2001) *Omitted variables in longitudinal data models* The Canadian Journal of Statistics vol. 29, No. 4, pp 573-595.

[10] Gaudry M., Lassarre S.(2000), *Structural Road Accidents Models: The international DRAG Family*, Pergamon Ed.

[11] Guttman L (1946) *Enlargement methods for computing the inverse matrix* Ann. Math. Statist. 17:336-343.

[12] Harville D.A. (1977) *Maximum Likelihood Approaches to Variance Component Estimation and to related Problems* J.A.S.A. vol. 72, n^0358, 320-340

[13] Hauer E. (1997), *Observasional before-after Studies in Road Safety: Estimating the effect of highway and traffic engineering measures on road safety*, Pergamon Edition.

[14] Le Nhu D., Sun L., Zidek J.V. (2001) *Spatial prediction and temporal backcasting for environmental fields having monotone data patterns* The Canadian Journal of Statistics vol. 29, No. 4, pages 529-554.

[15] Lee S-Y (1979) *Constrained estimation in covariance structure analysis* Biometrika 66, 3, 539-545.

[16] Neuenschwander B.E., Flury B.D. (1997) *A note on Silvey's (1959) theorem* Statistics & Probability Letter 36, 307-317.

[17] N'Guessan A. (1993), *Contribution à l'analyse statistique d'une mesure de sécurité routière*, Ph. D Thesis, Université de Lille 1. France.

[18] N'Guessan A. (2003a) *On a use of Schur complements for an inverse of a constrained accident data information matrix* PUB. IRMA, LILLE Vol. 60, N^0VI.

[19] N'Guessan A., Langrand C. (1993), *Sur la distribution asymptotique de certaines statistiques utilisées dans le domaine de la sécurité routière*, C.R. Acad. Sci, t.317, Série I, 401–404.

[20] N'Guessan A., Essai A., Langrand C. (2001) *Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière* Rev. Statistique Appliquée. XLIX(2), 83-100.

[21] Ouellette D.V. (1981) *Schur complements and statistics* Linear Algebra an its applications 36: 187-295.

[22] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. F. (1992) *Numerical Recipes in C: The Art of Scientific Computing* Second Ed. Cambrige University Press.

[23] Rivest L-P., Lévesque T. (2001) *Improved log-linear model estimators of abundance in capture-recapture experiments* The Canadian Journal of Statistics vol. 29, No. 4, pp 555-572.

[24] Sherman J., Morrison W. J. (1950) *Adjustement of an inverse matrix corresponding a change in one element of a given matrix* Ann. Math. Statist. 21:124-127.

[25] Silvey S.D. (1959) *The Langrangian multiplier test* Ann. Math. Statist. 30, 389-407.

[26] Tanner J.C. (1958), *A problem in the combination of accident frequencies*, Biometrika, 45, 331–342.

[27] Van Eeden C. (1996) *Estimation in Restricted Parameter Spaces - Some history and Some Recent developments* CWI Quarterly Vol. 9 (1 and 2) , 69 - 76

[28] Wald A. (1949) *Note on the consistency of the maximum likelihood estimate* Ann. Math. Statist. 20, 595-601.

[29] Woodburry M. A. (1950) *Inverting modified matrices* Memorandum Report 42,Statistical Research Group, Princeton, N.J.